



Noncoherent MIMO Multiple-Access Channels: A Joint Constellation Design

Downloaded from: <https://research.chalmers.se>, 2021-12-11 21:09 UTC

Citation for the original published paper (version of record):

Ngo, K., Yang, S., Guillaud, M. et al (2021)

Noncoherent MIMO Multiple-Access Channels: A Joint Constellation Design

2020 IEEE Information Theory Workshop, ITW 2020

<http://dx.doi.org/10.1109/ITW46852.2021.9457669>

N.B. When citing this work, cite the original published paper.

Noncoherent MIMO Multiple-Access Channels: A Joint Constellation Design

Khac-Hoang Ngo*, Sheng Yang†, Maxime Guillaud‡, Alexis Decurninge‡

*Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden

†Laboratory of Signals and Systems, CentraleSupélec, Paris-Saclay University, 91190 Gif-sur-Yvette, France

‡Mathematical and Algorithmic Sciences Laboratory, Huawei Technologies France, 92100 Boulogne-Billancourt, France

Emails: ngok@chalmers.se, sheng.yang@centralesupelec.fr, {maxime.guillaud, alexis.decurninge}@huawei.com

Abstract—We consider the joint constellation design problem for noncoherent multiple-input multiple-output multiple-access channels. By analyzing the noncoherent maximum-likelihood detection error, we propose novel design criteria so as to minimize the error probability. For any given set of constellation sizes, the proposed metrics can be optimized over the set of signal matrices. Based on these criteria, we propose a simple and efficient construction consisting in partitioning a single-user constellation. Numerical results show that our proposed metrics are meaningful, and can be used as objectives to generate constellations through numerical optimization that perform better, for the same transmission rate and power constraint, than a common pilot-based scheme and the constellations optimized with existing metrics.

I. INTRODUCTION

In multiple-input multiple-output (MIMO) communications, it is usually assumed that the channel state information (CSI) is known or estimated (typically using pilots and/or feedback), and then used for precoding at the transmitter and/or detection at the receiver. This is known as the *coherent* approach. On the other hand, in the *noncoherent* approach, the transmission and reception are designed without using the *a priori* knowledge of the CSI [1], [2], [3]. This paper studies the latter approach for the MIMO block-fading multiple-access channel (MAC), i.e., the channel is assumed to remain unchanged during each coherence block of length T and varies between blocks.

In the single-user case with isotropic Rayleigh fading, a noncoherent approach, so-called unitary space-time modulation (USTM), is to transmit $T \times M$ isotropically distributed and truncated unitary signal matrices, where M is the number of transmit antennas. Information is carried by the position of the signal matrix subspace in the Grassmann manifold $G(\mathbb{C}^T, M)$, defined as the space of M -dimensional subspaces in \mathbb{C}^T . This approach was shown to be within a vanishing gap from the high-SNR capacity if $T \geq N + \min\{M, N\}$ [1], [2], and a constant gap if $2M \leq T \leq M + N$ [3], where N is the number of receive antennas. There has been extensive research on the design of noncoherent constellations as a set of points on the Grassmann manifold, with a common criterion of maximizing the minimum pairwise chordal distance between the symbols [4].

In the multi-user case, a straightforward extension of the single-user coherent approach is to divide the coherence block into a training part in which orthogonal pilot sequences are sent to estimate the CSI for each user, and a data transmission part in which different users communicate in a nonorthogonal fashion. Although this approach achieves the optimal degree-of-freedom (DoF) region in the two-user SIMO MAC [5], its optimality

in terms of achievable rate and detection error probability remains unclear. An amplitude-based encoding scheme was proposed in [6], but the accompanying energy detector relies on a large number of receive antennas so that the average received power concentrates. Also with massive receive antenna array, some differential encoding schemes were proposed based on phase shift keying (PSK) [7] or quadrature amplitude modulation (QAM) [8]. In [9], we proposed a precoding-based multiple-access scheme allowing efficient detection, but offering no optimality guarantee.

In this work, we consider a K -user MIMO MAC with Rayleigh block fading with coherence time $T \geq 2$ where user k is equipped with M_k antennas and the receiver with N antennas. We aim to derive simple and effective joint constellation design criteria so as to minimize the joint maximum likelihood (ML) symbol detection error. If the users could cooperate, the system could be seen as a $(\sum_{k=1}^K M_k) \times N$ MIMO point-to-point channel. Thus the joint constellation can be treated as a Grassmannian constellation on $G(\mathbb{C}^T, \sum_{k=1}^K M_k)$, which leads to a design criterion mimicking the max-min chordal distance criterion. The error probability of the ML detector for the MIMO MAC was derived in [10]. With cooperating users, this analysis led to a design criterion similar to that for a single-user MIMO channel proposed in [11, Eq.(8)]. However, for non-cooperating users (as we consider here), using the same criterion would be suboptimal. The joint ML pairwise error exponent can be shown to converge to the Kullback-Leibler (KL) divergence between the output distributions conditioned on either of the transmitted symbols. Based on this analysis, a criterion consisting in maximizing the minimum KL divergence was proposed in [12], but was used only to optimize the transmit powers and the sub-constellation assignment.

In this paper, we analyze the worst-case pairwise error probability and introduce new constellation design metrics. The first proposed metric is the minimum expected pairwise log-likelihood ratio (PLLR) over the joint constellation. This coincides with the minimum KL divergence metric in the massive MIMO regime. From the dominant term of the expected PLLR at high SNR, we obtain further simplified metrics. We also propose an alternating optimization consisting in iteratively optimizing one user at a time to simplify the optimization. For any given set of constellation sizes, the proposed metrics can be optimized over the set of signal matrices. Based on our metrics, we propose a simple construction consisting in partitioning a single-user constellation. Assuming unitary

space-time modulation, we numerically optimize the metrics to generate joint constellations, and compare with a pilot-based constellation and constellations optimized with existing metrics, including the one in [11]. Numerical results show that our proposed metrics are meaningful and effective, and the resulting constellations outperform the aforementioned baselines.

The remainder of the paper is organized as follows. In Sec. II, we present the system model and formulate the problem. In Sec. III, we analyze the detection error probability and derive the design metrics, based on which we propose a partitioning-based constructions in Sec. IV. We present the numerical results in Sec. V and conclude the paper in Sec. VI. The proofs can be found in the extended version [13].

Notation: Random quantities are denoted with non-italic letters with sans-serif fonts, e.g., a scalar \mathbf{x} , a vector \mathbf{v} , and a matrix \mathbf{M} . Deterministic quantities are denoted with italic letters, e.g., a scalar x , a vector v , and a matrix M . The Euclidean norm is denoted by $\|\cdot\|$ and the Frobenius norm by $\|\cdot\|_F$. The trace, transpose, conjugate, and conjugate transpose of \mathbf{M} are respectively $\text{tr}(\mathbf{M})$, \mathbf{M}^\top , \mathbf{M}^* , and \mathbf{M}^H . $[n] := \{1, 2, \dots, n\}$. $\mathbb{1}\{\cdot\}$ is the indicator function.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a MIMO MAC consisting of a receiver equipped with N antennas and K users, user k with M_k antennas, $k \in [K]$. The channel is assumed to be flat and block fading with equal-length and synchronous (across the users) coherence intervals of length $T \geq 2$. That is, the channel matrix $\mathbf{H}_k \in \mathbb{C}^{N \times M_k}$ of user k remains constant within each coherence block of T channel uses and changes between blocks. Furthermore, the *distribution* of \mathbf{H}_k is assumed to be known, but its *realizations* are unknown to both the users and the receiver. We consider independent and identically distributed (i.i.d.) Rayleigh fading, namely, the rows of $\mathbf{H} := [\mathbf{H}_1 \ \mathbf{H}_2 \ \dots \ \mathbf{H}_K]$ are independent and follow $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{M_{\text{tot}}})$ where $M_{\text{tot}} := \sum_{k=1}^K M_k$. Within a coherence block t , each user k sends a signal matrix symbol $\mathbf{X}_k \in \mathbb{C}^{T \times M_k}$, and the receiver observes

$$\mathbf{Y}[t] = \sum_{k=1}^K \mathbf{X}_k[t] \mathbf{H}_k^\top[t] + \mathbf{Z}[t], \quad t = 1, 2, \dots \quad (1)$$

where the additive noise $\mathbf{Z} \in \mathbb{C}^{T \times N}$ has i.i.d. $\mathcal{CN}(0, 1)$ entries independent of $\{\mathbf{H}_k\}$. We consider the power constraint $\frac{1}{n} \sum_{t=1}^n \|\mathbf{X}_k[t]\|_F^2 \leq PT, k \in [K]$ where n is the number of blocks spanned by a codeword. Thus, P is also an upper bound of the per-user SNR at each receive antenna.

We assume that \mathbf{X}_k takes value from a *finite constellation* \mathcal{X}_k of fixed size $|\mathcal{X}_k| = 2^{R_k T}$ with equally likely symbols, where R_k (bits/channel use) is the transmission rate. To satisfy the power constraint, we assume that $\frac{1}{|\mathcal{X}_k|} \sum_{\mathbf{X}_k \in \mathcal{X}_k} \|\mathbf{X}_k\|_F^2 =: P_k T, k \in [K]$, and $\max_k P_k = P$. Let us focus on one block, omit the block index, and rewrite (1) as

$$\mathbf{Y} = [\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_K] [\mathbf{H}_1 \ \mathbf{H}_2 \ \dots \ \mathbf{H}_K]^\top + \mathbf{Z} = \mathbf{X} \mathbf{H}^\top + \mathbf{Z},$$

where the concatenated signal matrix $\mathbf{X} := [\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_K]$ takes value from $\mathcal{X} := \{[\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_K] : \mathbf{X}_k \in \mathcal{X}_k\} = \prod_{k=1}^K \mathcal{X}_k$. Our goal is to derive the desirable properties of the

set tuple $(\mathcal{X}_1, \dots, \mathcal{X}_K)$ for a given rate tuple (R_1, \dots, R_K) to achieve low *symbol detection error probability*.

The likelihood function $p_{\mathbf{Y}|\mathbf{X}}$ is given by

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X}) = \frac{\exp(-\text{tr}(\mathbf{Y}^H(\mathbf{I}_T + \mathbf{X}\mathbf{X}^H)^{-1}\mathbf{Y}))}{\pi^{NT} \det^N(\mathbf{I}_T + \mathbf{X}\mathbf{X}^H)}. \quad (2)$$

Therefore, given the received symbol $\mathbf{Y} = \mathbf{Y}$, the joint-user ML symbol detector is

$$\Xi(\mathbf{Y}) = \arg \max_{\mathbf{X} \in \mathcal{X}} \left(-\text{tr}((\mathbf{I}_T + \mathbf{X}\mathbf{X}^H)^{-1}\mathbf{Y}\mathbf{Y}^H) - N \ln \det(\mathbf{I}_T + \mathbf{X}\mathbf{X}^H) \right). \quad (3)$$

We aim to design \mathcal{X} so as to minimize the ML detection error $P_e(\mathcal{X}) = \mathbb{P}(\Xi(\mathbf{Y}) \neq \mathbf{X})$, i.e.,

$$\mathcal{X}^* = \arg \max_{\mathcal{X}} P_e(\mathcal{X}).$$

Since $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X})$ depends on \mathbf{X} only through $\mathbf{X}\mathbf{X}^H$, the following proposition is straightforward.

Proposition 1 (Identifiability condition). *For $P_e(\mathcal{X})$ to vanish at high SNR, the joint constellation \mathcal{X} must satisfy $\mathbf{X}\mathbf{X}^H \neq \mathbf{X}'\mathbf{X}'^H$ for any pair of distinct symbols \mathbf{X} and \mathbf{X}' in \mathcal{X} .*

In the next section, we analyze the error probability and derive more specific design criteria.

III. CONSTELLATION DESIGN CRITERIA

With \mathbf{X} uniformly distributed in \mathcal{X} , $P_e(\mathcal{X})$ can be written as

$$P_e(\mathcal{X}) = \frac{1}{|\mathcal{X}|} \sum_{\mathbf{X} \in \mathcal{X}} \mathbb{P}(\Xi(\mathbf{Y}) \neq \mathbf{X} | \mathbf{X} = \mathbf{X}). \quad (4)$$

We denote the pairwise error event as $\{\mathbf{X} \rightarrow \mathbf{X}'\} := \{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X}) \leq p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X}') | \mathbf{X} = \mathbf{X}\}$. Since the ML detection error event is the union of the pairwise error events, after some simple manipulations, we have the following bounds on $P_e(\mathcal{X})$:

$$\frac{1}{|\mathcal{X}|} \max_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \mathbb{P}(\mathbf{X} \rightarrow \mathbf{X}') \leq P_e(\mathcal{X}) \leq (|\mathcal{X}| - 1) \max_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \mathbb{P}(\mathbf{X} \rightarrow \mathbf{X}').$$

We see that for a given constellation size $|\mathcal{X}|$, $P_e(\mathcal{X})$ vanishes if and only if the *worst-case pairwise error probability (PEP)*, $\max_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \mathbb{P}(\mathbf{X} \rightarrow \mathbf{X}')$, vanishes. Therefore, our goal from now on is to minimize the worst-case PEP. Let us rewrite the PEP as $\mathbb{P}(\mathbf{X} \rightarrow \mathbf{X}') = \mathbb{P}(\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}') \leq 0)$ with the PLLR $\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}') := \ln \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X})}{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X}')}$. Using (2), we obtain

$$\begin{aligned} \mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}') &= N \ln \frac{\det(\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)}{\det(\mathbf{I}_T + \mathbf{X}\mathbf{X}^H)} \\ &\quad - \text{tr}(((\mathbf{I}_T + \mathbf{X}\mathbf{X}^H)^{-1} - (\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)^{-1})\mathbf{Y}\mathbf{Y}^H). \end{aligned} \quad (5)$$

Although a closed-form expression of the PEP can be derived following [10, Prop. 1], it does not bring clear insights into the constellation design. A high-SNR asymptotic expression of the PEP was given in [10, Prop. 3], but is also hard to exploit. An alternative approach is to treat the joint constellation as a constellation of an $M_{\text{tot}} \times N$ MIMO point-to-point channel,

and adopt the single-user max-min chordal distance criterion $\mathcal{X}^* = \arg \max_{\mathcal{X}} \min_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \text{tr} \left(\mathbf{I} - \frac{\mathbf{X}^H \mathbf{X}' \mathbf{X}'^H \mathbf{X}}{\|\mathbf{X}\|_F^2 \|\mathbf{X}'\|_F^2} \right)$, or equivalently,

$$\mathcal{X}^* = \arg \min_{\mathcal{X}} \max_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \text{tr} \left(\frac{\mathbf{X}^H \mathbf{X}' \mathbf{X}'^H \mathbf{X}}{\|\mathbf{X}\|_F^2 \|\mathbf{X}'\|_F^2} \right). \quad (6)$$

Another criterion inspired by the single-user criterion proposed in [11, Eq.(8)] is

$$\mathcal{X}^* = \arg \min_{\mathcal{X}} \sum_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \det^{-N} \left(\mathbf{I} - \frac{\mathbf{X}^H \mathbf{X}' \mathbf{X}'^H \mathbf{X}}{\|\mathbf{X}\|_F^2 \|\mathbf{X}'\|_F^2} \right). \quad (7)$$

Let us denote $m_1(\mathcal{X}) := \max_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \text{tr} \left(\frac{\mathbf{X}^H \mathbf{X}' \mathbf{X}'^H \mathbf{X}}{\|\mathbf{X}\|_F^2 \|\mathbf{X}'\|_F^2} \right)$ and $m_2(\mathcal{X}) := \ln \sum_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \det^{-N} \left(\mathbf{I} - \frac{\mathbf{X}^H \mathbf{X}' \mathbf{X}'^H \mathbf{X}}{\|\mathbf{X}\|_F^2 \|\mathbf{X}'\|_F^2} \right)$, and refer to (6) and (7) respectively as Min- m_1 and Min- m_2 . Next, we present our approach and derive new design criteria.

A. Design Criteria

We resort to the following bound on the PEP

$$\mathbb{P}(\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}') \leq 0) \leq \frac{\text{Var}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')] }{\text{Var}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')] + \mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]^2} \quad (8)$$

which follows from Cantelli's inequality.¹ Note that the upper bound decreases when $\frac{\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]^2}{\text{Var}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]}$ increases. We choose to relax the problem into maximizing the expected PLLR $\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]$. Although maximizing $\frac{\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]^2}{\text{Var}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]}$ and maximizing $\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]$ are equivalent only when $\text{Var}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]$ is constant over different symbol pairs, the relaxation makes the problem tractable.

We further justify our choice by pointing out the connection to the following hypothesis testing problem. Let us consider two hypotheses: $H_0 : \{\mathbf{y}_i\}_{i=1}^N \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_T + \mathbf{X}\mathbf{X}^H)$ and $H_1 : \{\mathbf{y}_i\}_{i=1}^N \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)$ where $\{\mathbf{y}_i\}_{i=1}^N$ are realizations of N columns of \mathbf{Y} . Then, the PEP $\mathbb{P}(\mathbf{X} \rightarrow \mathbf{X}')$ can be seen as the type-1 error probability of the optimal likelihood ratio test. From (8) and the fact that $\frac{\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]}{\text{Var}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')] } \rightarrow \infty$ as $N \rightarrow \infty$, we have that $\mathbb{P}(\mathbf{X} \rightarrow \mathbf{X}') \rightarrow 0$ as $N \rightarrow \infty$ for any constellation satisfying the identifiability condition in Prop. 1. (A proof is given in [13, Appendix C].) Switching the symbols' roles, we deduce that $\mathbb{P}(\mathbf{X}' \rightarrow \mathbf{X}) \leq \epsilon \in (0, 1/2)$ for N large enough. Then, it follows from the Chernoff-Stein Lemma that

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}(\mathbf{X} \rightarrow \mathbf{X}') \\ &= -D(\mathcal{CN}(\mathbf{0}, \mathbf{I}_T + \mathbf{X}\mathbf{X}^H) \parallel \mathcal{CN}(\mathbf{0}, \mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)) \\ &= -\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')], \end{aligned}$$

where $D(\cdot \parallel \cdot)$ is the KL divergence. Thus, maximizing $\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]$ maximizes the pairwise error exponent when $N \rightarrow \infty$.

Therefore, letting $e_{\min}(\mathcal{X}) := \frac{1}{N} \min_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]$, we consider the following design criterion and refer to it as Max- e_{\min} :

$$\mathcal{X}^* = \arg \max_{\mathcal{X}} e_{\min}(\mathcal{X}) \quad (9)$$

¹Cantelli's inequality states that $\mathbb{P}(x - \mu \leq \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}$ for a real-valued random variable x with mean μ and variance σ^2 , and $\lambda < 0$. Applying this with $x = \mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')$ and $\lambda = -\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')]$, we obtain (8).

where it follows from (5) and $\mathbb{E}[\mathbf{Y}\mathbf{Y}^H] = N(\mathbf{I}_T + \mathbf{X}\mathbf{X}^H)$ that $\mathbb{E}[\mathbf{L}(\mathbf{X} \rightarrow \mathbf{X}')] = N \ln \frac{\det(\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)}{\det(\mathbf{I}_T + \mathbf{X}\mathbf{X}^H)} - N + N \text{tr}((\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)^{-1}) + N \text{tr}((\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)^{-1} \mathbf{X}\mathbf{X}^H)$. (10)

Lemma 1. Let \mathbf{X} and \mathbf{X}' be s.t. $\|\mathbf{X}\|_F^2 = \Theta(P)$ and $\|\mathbf{X}'\|_F^2 = \Theta(P)$ as $P \rightarrow \infty$. We have $\text{tr}((\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)^{-1}) = O(1)$; $\ln \frac{\det(\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)}{\det(\mathbf{I}_T + \mathbf{X}\mathbf{X}^H)} = O(1)$ if $\text{Span}(\mathbf{X}) = \text{Span}(\mathbf{X}')$ and $O(\ln P)$ otherwise. Furthermore, $\text{tr}((\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)^{-1} \mathbf{X}\mathbf{X}^H) = O(1)$ if $\text{Span}(\mathbf{X}) = \text{Span}(\mathbf{X}')$ and $O(P)$ otherwise.

Proof. See [13, Appendix D]. \square

We see that $d(\mathbf{X} \rightarrow \mathbf{X}') := \text{tr}((\mathbf{I}_T + \mathbf{X}'\mathbf{X}'^H)^{-1} \mathbf{X}\mathbf{X}^H)$ is the only term in (10) that can scale up linearly with P . Letting $d_{\min}(\mathcal{X}) := \min_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} d(\mathbf{X} \rightarrow \mathbf{X}')$, we have the following design criterion, referred to as Max- d_{\min} ,

$$\mathcal{X}^* = \arg \max_{\mathcal{X}} d_{\min}(\mathcal{X}). \quad (11)$$

Hereafter, we assume for simplicity that all users have the same number of antennas, i.e. $M_1 = \dots = M_K = M$, although the general case follows in a straightforward manner.

B. The Single-User Case

In the single-user case with M transmit antennas, it is known that USTM is optimal, or near optimal, at the high SNR regime [2], [3]. We consider this approach and let $\mathbf{X}^H \mathbf{X} = \frac{PT}{M} \mathbf{I}_M, \forall \mathbf{X} \in \mathcal{X}$. It follows that $d(\mathbf{X} \rightarrow \mathbf{X}') = PT \left(1 - \alpha_{P,T,M} \frac{\|\mathbf{X}^H \mathbf{X}'\|_F^2}{(PT)^2} \right)$, where $\alpha_{P,T,M} := \left(\frac{1}{PT} + \frac{1}{M} \right)^{-1}$. Therefore, the design criterion (11) is equivalent to $\mathcal{X} = \arg \min_{\mathcal{X}} \max_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \|\mathbf{X}^H \mathbf{X}'\|_F^2$. This coincides with the commonly used max-min pairwise chordal distance criterion [4].

C. The K -User Case

In the K -user case, we have the following proposition.

Proposition 2. It holds that

$$\min_{k \in [K]} d_k(\mathcal{X}) \leq d_{\min}(\mathcal{X}) \leq \min_{k \in [K]} d_k(\mathcal{X}) + (K-1)M, \quad (12)$$

where $d_k(\mathcal{X}) := \min_{\substack{\mathbf{x}_k \neq \mathbf{x}'_k \in \mathcal{X}_k \\ \mathbf{x}_j \in \mathcal{X}_j, j \neq k}} \text{tr} \left(\mathbf{x}_k^H (\mathbf{I}_T + \mathbf{x}'_k \mathbf{x}'_k{}^H + \sum_{j \neq k} \mathbf{x}_j \mathbf{x}_j{}^H)^{-1} \mathbf{x}_k \right)$.

Proof. See [13, Appendix E]. \square

Proposition 2 says that $d_{\min}(\mathcal{X})$ is within a constant gap to $\min_{k \in [K]} d_k(\mathcal{X})$, and thus $d_{\min}(\mathcal{X})$ scales linearly with P when P is large if and only if $\min_{k \in [K]} d_k(\mathcal{X})$ does so. Based on this observation, we propose the following design criterion

$$\mathcal{X}^* = \arg \max_{\mathcal{X}} \min_{k \in [K]} d_k(\mathcal{X}). \quad (13)$$

To simplify the constellation optimization, we propose an *alternating optimization* approach as follows. First $\{\mathcal{X}_k\}_{k=1}^K$ are initialized. Then, for $k \in [K]$, we iteratively optimize \mathcal{X}_k by $\mathcal{X}_k^* = \arg \max_{\mathcal{X}_k} m(\mathcal{X})$ for fixed $\{\mathcal{X}_l\}_{l \neq k}$ in a round robin manner, where $m(\mathcal{X})$ is the considered metric. At each iteration, it has fewer variables to optimize than directly solving (9), (11), or (13). However, the solution may not converge, and when it does, it is not guaranteed to converge to the optimum.

IV. CONSTELLATION DESIGN BASED ON PARTITIONING

In this section, we consider the symmetrical power case $P_k = P$, $k \in [K]$. This is a reasonable assumption if the rates are symmetric. To reduce the solution space, we make the suboptimal assumption that the individual constellations \mathcal{X}_k follow from USTM, i.e., they contain scaled-truncated-unitary-matrix symbols. From a practical perspective, this is desirable since the constellation is oblivious to the presence of the other users and USTM is high-SNR optimal, or near optimal, for the single-user channel. Nevertheless, there must be constraints between the symbols of different users. For instance, if the constellations are s.t. $\mathbf{X}_1 = \mathbf{X}_2$ can occur, then $d_k(\mathcal{X})$ is upper-bounded by a constant for any k and any P . This can be developed in a formal way as follows. By removing the terms inside the inverse in $d_k(\mathcal{X})$, we obtain an upper bound:

$$d_k(\mathcal{X}) \leq \min \left\{ \min_{\mathbf{X}_k \neq \mathbf{X}'_k \in \mathcal{X}_k} \text{tr}(\mathbf{X}_k^H (\mathbf{I}_T + \mathbf{X}'_k \mathbf{X}'_k^H)^{-1} \mathbf{X}_k), \right. \\ \left. \min_{\mathbf{X}_k \in \mathcal{X}_k, \mathbf{X}_l \in \mathcal{X}_l, l \neq k} \text{tr}(\mathbf{X}_k^H (\mathbf{I}_T + \mathbf{X}_l \mathbf{X}_l^H)^{-1} \mathbf{X}_k) \right\}. \quad (14)$$

For $d_k(\mathcal{X})$ to be large, the upper bound must be large. This is made precise in the next proposition.

Proposition 3 (Necessary condition). *Let $\{\mathcal{X}_k\}_{k=1}^K$ be s.t. $\mathbf{X}_k^H \mathbf{X}_k = \frac{PT}{M} \mathbf{I}_M$, $\forall \mathbf{X}_k \in \mathcal{X}_k$, $k \in [K]$. If the following lower bound on the d -values holds for some $c \in [0, 1/M]$*

$$\min_{k \in [K]} d_k(\mathcal{X}) \geq PT(1 - \alpha_{P,T,M} c),$$

where $\alpha_{P,T,M} := (\frac{1}{PT} + \frac{1}{M})^{-1}$, then we must have

$$\frac{1}{(PT)^2} \max \left\{ \max_{\mathbf{X}_k \neq \mathbf{X}'_k \in \mathcal{X}_k, k \in [K]} \|\mathbf{X}'_k{}^H \mathbf{X}_k\|_F^2, \right. \\ \left. \max_{\mathbf{X}_k \in \mathcal{X}_k, \mathbf{X}_l \in \mathcal{X}_l, k \neq l \in [K]} \|\mathbf{X}_k^H \mathbf{X}_l\|_F^2 \right\} \leq c. \quad (15)$$

Proof. The proof follows the same steps as in the single-user case, applying to the upper bound (14). \square

The above proposition shows that symbol pairs from different users should fulfill similar distance criteria as symbol pairs from the same user when it comes to identifiability conditions. However, it is unclear whether (15) alone is enough to guarantee a large value of $d_{\min}(\mathcal{X})$. In the following, we shall show that these conditions are indeed sufficient if c is small.

Proposition 4 (Sufficient condition). *Let $\{\mathcal{X}_k\}_{k=1}^K$ be s.t. $\mathbf{X}_k^H \mathbf{X}_k = \frac{PT}{M} \mathbf{I}_M$, $\forall \mathbf{X}_k \in \mathcal{X}_k$, $k \in [K]$. If*

$$\frac{1}{(PT)^2} \max \left\{ \max_{\mathbf{X}_k \neq \mathbf{X}'_k \in \mathcal{X}_k, k \in [K]} \|\mathbf{X}'_k{}^H \mathbf{X}_k\|_F^2, \right. \\ \left. \max_{\mathbf{X}_k \in \mathcal{X}_k, \mathbf{X}_l \in \mathcal{X}_l, k \neq l \in [K]} \|\mathbf{X}_k^H \mathbf{X}_l\|_F^2 \right\} \leq c$$

for some $c \in [0, 1/M]$, then we have

$$\min_{k \in [K]} d_k(\mathcal{X}) \geq PT \left(1 - K \left(\frac{1}{PT} + \frac{1}{M} - \sqrt{\frac{K(K-1)c}{2^{1\{K=2\}}}} \right)^{-1} c \right).$$

Proof. See [13, Appendix F]. \square

The two propositions above motivate the following simplified design criterion

$$\mathcal{X}^* = \arg \min_{\mathcal{X}} \max \left\{ \max_{\mathbf{X}_k \neq \mathbf{X}'_k \in \mathcal{X}_k, k \in [K]} \|\mathbf{X}'_k{}^H \mathbf{X}_k\|_F^2, \right. \\ \left. \max_{\mathbf{X}_k \in \mathcal{X}_k, \mathbf{X}_l \in \mathcal{X}_l, k \neq l \in [K]} \|\mathbf{X}_k^H \mathbf{X}_l\|_F^2 \right\}. \quad (16)$$

Based on (16), we propose a simple construction as follows. Let \mathcal{X}_{SU} be a single-user constellation and let $c := \frac{1}{(PT)^2} \max_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}_{\text{SU}}} \|\mathbf{X}'^H \mathbf{X}\|_F^2 \in [0, \frac{1}{M}]$. We can generate $\{\mathcal{X}_k\}_{k=1}^K$ by partitioning \mathcal{X}_{SU} into K disjoint subsets. Then, from (12) and Proposition 4, we can guarantee

$$d_{\min}(\mathcal{X}) \geq PT \left(1 - K \left(\frac{1}{PT} + \frac{1}{M} - \sqrt{\frac{K(K-1)c}{2^{1\{K=2\}}}} \right)^{-1} c \right). \quad (17)$$

With such a construction, the joint constellation design problem becomes essentially an individual constellation design problem. A random partition suffices to guarantee (17), although one can smartly partition the set \mathcal{X}_{SU} to improve over (17). The optimal partition problem is equivalent to a min-max graph partitioning. Note that for the right-hand side of (17) to scale linearly with P , we must have that $c < \left[\left(\frac{1}{KPT} + \frac{1}{KM} + \frac{K-1}{4K2^{1\{K=2\}}} \right)^{\frac{1}{2}} - \sqrt{\frac{K-1}{4K2^{1\{K=2\}}}} \right]^2$, which limits the size of the mother constellation \mathcal{X}_{SU} .

V. NUMERICAL RESULTS

In this section, we consider the single transmit antenna case ($M_k = 1$, $k \in [K]$) and assume USTM, i.e., $\mathcal{X}_k = \{\sqrt{P_k T} \mathbf{x}_k^{(i)}\}_{i=1}^{|\mathcal{X}_k|}$ with each $\mathbf{x}_k^{(i)}$ being a unit-norm vector. We will compare our design in this paper with our precoding design in [9] (with Precoder Type II therein), and the constellations optimized with the criteria Min- m_1 (6) and Min- m_2 (7) in terms of joint-ML symbol error rate (SER) (4). We also consider a pilot-based joint constellation corresponding to the scenario where K users transmit orthogonal pilots, followed by spatially multiplexed QAM symbols. The receiver uses either an ML detector (3) or a linear minimum-mean-square-error (MMSE) detector consisting of MMSE channel estimation, MMSE equalization, and component-wise QAM demapper.

Numerical Optimization: We solve numerically Max- e_{\min} (9), Max- d_{\min} (11) and the alternating optimization for given powers $\{P_k\}$. In general, we want to solve the manifold-constrained optimization $\max_{\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_K} \min_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} f(\mathbf{X}, \mathbf{X}')$, where $f(\mathbf{X}, \mathbf{X}')$ is customized according to the considered criterion. Note that the objective function is not smooth because of the min. To smooth it, we use the approximation $\max_i x_i \approx \epsilon \ln \sum_i \exp(x_i/\epsilon)$ with a small ϵ and obtain $\min_{\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_K} \epsilon \ln \sum_{\mathbf{X} \neq \mathbf{X}' \in \mathcal{X}} \exp(-\frac{f(\mathbf{X}, \mathbf{X}')}{\epsilon})$. (For other smoothing techniques, see, e.g., [14].) We compute the Riemannian gradient of the new objective function and resort to the Manopt toolbox [15] to solve the optimization by conjugate gradient descent on the manifold. The criteria Min- m_1 and Min- m_2 are solved numerically in a similar manner.

We show next some numerical results for the two-user case with the symmetrical rate $R_1 = R_2 = B/T$ and equal power $P_1 = P_2 = P$. (Thus P is identified with the SNR.) We

optimize the joint constellation at $P = 30$ dB, even when the performance is benchmarked at other SNR values. In all figures, the legends representing our schemes are in bold face.

First, in Fig. 1, we plot the joint SER (4) of the considered schemes for $T = 5$, $B = 5$, and $N = 4$. We observe that the constellations optimized with our metrics $e_{\min}(\mathcal{X})$ (9) and $d_{\min}(\mathcal{X})$ (11) achieve similar performance and are the best among the schemes with the same rate pair. The performance of the alternatively optimized constellation is slightly inferior and better than the pilot-based scheme. The partitioning design (with random partition) and the precoding design have similar performance. The constellations optimized with the Min- m_1 and Min- m_2 criteria, especially the latter, perform worse than that with Max- e_{\min} and Max- d_{\min} .

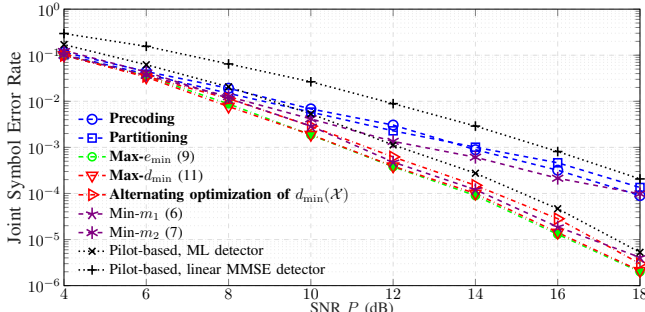


Fig. 1. The joint SER of the proposed constellations compared to the baselines for $T = 5$, $K = 2$, $B = 5$, and $N = 4$.

To assess the design metrics, we show the values of our metrics $d_{\min}(\mathcal{X})$ and $e_{\min}(\mathcal{X})$, and the other metrics $m_1(\mathcal{X})$ and $m_2(\mathcal{X})$ for these constellations in the same setting in Fig. 2. In Fig. 2(a), $d_{\min}(\mathcal{X})$ is very close to $e_{\min}(\mathcal{X})$ for $\text{SNR} \geq 20$ dB. The constellations with low joint-ML SER in Fig. 1 exhibit a large value of these metrics. This confirms that our proposed metrics are meaningful for constellation design and evaluation. From Fig. 2(b) and Fig. 2(c), we see that the relative order of the constellations in terms of the value of $m_1(\mathcal{X})$ and $m_2(\mathcal{X})$ is unrevealing about the SER performance in Fig. 1. In particular, although the constellation optimized with Min- m_1 (6) also achieves a low joint-ML SER (and a high value of our metrics), this is not true for other constellations. For example, the partitioning design has a low m_1 -value but a high SER; the pilot-based constellation has a much lower m_2 -value than the Max- e_{\min} design, but has a higher SER.

VI. CONCLUSION

This work addresses the joint constellation design for noncoherent MIMO MAC in Rayleigh block fading. By analyzing the joint detection error, we have derived closed-form metrics which are effective for designing joint constellations that achieve a low error rate. In this work, we have focused on the optimality with respect to the joint maximum-likelihood detector, which has high complexity in general. It would be interesting, as in the single-user case, to construct structured joint constellations that allow for low-complexity detection.

REFERENCES

[1] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.

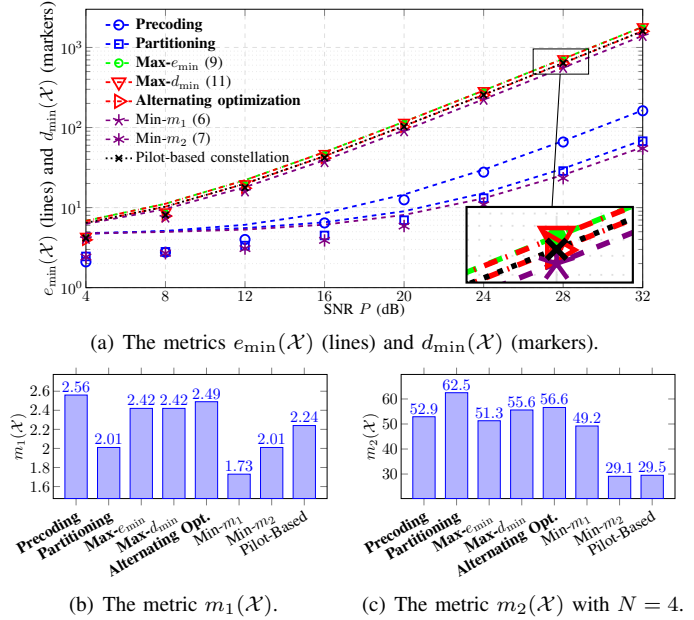


Fig. 2. The value of the metrics $e_{\min}(\mathcal{X})$, $d_{\min}(\mathcal{X})$, $m_1(\mathcal{X})$ and $m_2(\mathcal{X})$ for the considered constellations for $T = 5$ and $B = 5$.

[2] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[3] W. Yang, G. Durisi, and E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 117–132, Feb. 2013.

[4] R. H. Gohary and H. Yanikomeroglu, "Noncoherent MIMO signaling for block-fading channels: Approaches and challenges," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 80–88, Mar. 2019.

[5] K.-H. Ngo, S. Yang, and M. Guillaud, "The optimal DoF region for the two-user non-coherent SIMO multiple-access channel," in *IEEE Information Theory Workshop (ITW)*, Guangzhou, China, Nov. 2018.

[6] A. Manolakis, M. Chowdhury, and A. Goldsmith, "Energy-based modulation for noncoherent massive SIMO systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7831–7846, Nov 2016.

[7] V. M. Baeza, A. G. Armada, W. Zhang, M. El-Hajjar, and L. Hanzo, "A noncoherent multiuser large-scale SIMO system relying on M-ary DPSK and BICM-ID," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1809–1814, Feb 2018.

[8] D. Kong, X. Xia, and T. Jiang, "A differential QAM detection in uplink massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6371–6383, Sep. 2016.

[9] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "A multiple access scheme for non-coherent SIMO communications," in *52nd Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Oct. 2018, pp. 1846–1850.

[10] M. Brehler and M. K. Varanasi, "Noncoherent multiuser space-time communications: Optimum receivers and signal design," in *Proceedings of the 35th Annual Conference Information Science System*. The Johns Hopkins University, Mar. 2001, pp. 379–383.

[11] M. L. McCloud, M. Brehler, and M. K. Varanasi, "Signal design and convolutional coding for noncoherent space-time communication on the block-Rayleigh-fading channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1186–1194, May 2002.

[12] H. Chen, Z. Dong, J.-K. Zhang, and B. Vucetic, "Design of non-orthogonal and noncoherent massive MIMO for scalable URLLC beyond 5G," *arXiv preprint arXiv:2001.10728*, 2020.

[13] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Joint constellation design for non-coherent MIMO multiple-access channels," *submitted to IEEE Trans. Inf. Theory*, 2020, arXiv preprint arXiv:2009.11548.

[14] A. Medra and T. N. Davidson, "Flexible codebook design for limited feedback systems via sequential smooth optimization on the grassmannian manifold," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1305–1318, Jan. 2014.

[15] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1455–1459, Jan. 2014.